

AMENDMENTS TO CLAIMS:

The listing of claims will replace all prior versions, and listings, of claims in the application:

LISTING OF CLAIMS:

1-6 (Cancelled)

7. (Previously Presented) A document categorizing method for categorizing a plurality of documents in an electronic system into a plurality of clusters according to semantic similarity, said method being characterized in that:

after categorizing said plurality of documents into a plurality of clusters according to semantic similarity, a cluster merging process is performed such that relations among clusters of said plurality of clusters are evaluated on the basis of documents included in the respective clusters, and two or more clusters having a degree of relation equal to or higher than a first predetermined value are combined together;

wherein the cluster names of respective clusters merged together are displayed such that when said degree of relation among said clusters is higher than a second predetermined value higher than said first predetermined value, said cluster names are displayed in a first listing format, and when said degree of relation among said clusters is lower than said second predetermined value and higher than said first predetermined value, said cluster names are displayed in a second listing format.

8. (Previously Presented) A document categorizing method according to Claim 7, wherein when said cluster names are displayed in said first listing format, said cluster names of the respective clusters are displayed successively in a single horizontal line or are displayed successively in different lines, and when said cluster names are displayed in said second listing format, a delimiter is inserted between adjacent cluster names of the respective clusters.

9. (Previously Presented) A document categorizing method according to Claim 7, wherein when a first cluster includes a second cluster therein, the name of

said second cluster included in said first cluster is enclosed within brackets and placed after the name of said first cluster.

10. (Cancelled)

11. (Presently Amended) A document categorizing apparatus for categorizing a plurality of electronic documents into a plurality of clusters according to semantic similarity, said apparatus comprising:

a clustering unit for categorizing a plurality of documents into a plurality of clusters in accordance with semantic similarity,

a cluster merging unit for evaluating the relation among the plurality of clusters created by said clustering unit on the basis of the documents included in the respective clusters and then combines two or more clusters having a degree of relation equal to or higher than a first predetermined value;

an output display for displaying the cluster names of respective clusters merged together such that when said degree of relation among said clusters is higher than a second predetermined value higher than said first predetermined value, said cluster names are displayed in a first listing format, and when said degree of relation among said clusters is lower than said second predetermined value and higher than said first predetermined value, said cluster names are displayed in a second listing format.

12. (Cancelled)

13. (Previously Presented) A storage medium on which a document categorizing program for categorizing a plurality of documents into a plurality of clusters according to semantic similarity is stored, said document categorizing program comprising:

a clustering step for categorizing a plurality of documents into a plurality of clusters in accordance with semantic similarity;

a cluster merging step in which the degrees of relation among clusters of said plurality of clusters obtained in said clustering step are evaluated on the basis of documents included in the respective clusters, and two or more clusters

having a degree of relation equal to or higher than a first predetermined value are combined together; and

a step for outputting said cluster-merging-process information, wherein the cluster names of respective clusters merged together are displayed such that when said degree of relation among said clusters is higher than a second predetermined value higher than said first predetermined value, said cluster names are displayed in a first listing format, and when said degree of relation among said clusters is lower than said second predetermined value and higher than said first predetermined value, said cluster names are displayed in a second listing format.

14. (Cancelled)

15. (Previously Presented) The document categorizing method of Claim 7, wherein said first listing format is an AND listing format and said second listing format is an OR listing format.

16 -19 (Cancelled)

20. (Previously Presented) The document categorizing apparatus of claim 11, wherein said first listing format is an AND listing format and said second listing format is an OR listing format.

21. (Previously Presented) The document categorizing apparatus of claim 11, wherein when said cluster names are displayed in said first listing format, said cluster names of the respective clusters are displayed successively in a single horizontal line or are displayed successively in different lines, and when said cluster names are displayed in said second listing format, a delimiter is inserted between adjacent cluster names of the respective clusters.

22. (Previously Presented) The storage medium of claim 13, wherein said first listing format is an AND listing format and said second listing format is an OR listing format.

23. (Previously Presented) The storage medium of claim 13, wherein when said cluster names are displayed in said first listing format, said cluster names of the

respective clusters are displayed successively in a single horizontal line or are displayed successively in different lines, and when said cluster names are displayed in said second listing format, a delimiter is inserted between adjacent cluster names of the respective clusters.

24. (New) A document categorizing method for categorizing a plurality of documents in an electronic system according to semantic similarity, said method comprising:

obtaining a plurality of clusters of documents, each cluster having a distinctive name;

evaluating a degree of relation between at least two clusters by evaluating the similarity between the evaluated clusters based on the documents included in the respective evaluated clusters;

merging the evaluated clusters into a new combined cluster when their degree of relation is determined to be not less than a predetermined first value; and

assigning a new name to said new combined cluster based on the degree of relation between its constituent evaluated clusters;

wherein:

if the degree of relation of said constituent evaluated clusters is less than a second predetermined value, which is greater than said first predetermined value, the new name assigned to said new combined cluster conforms to a first naming convention indicative of a degree of relation between said first and second predetermined values; and

if the degree of relation of said constituent evaluated clusters is not less than said second predetermined value, the new name assigned to said new combined cluster conforms to a second naming convention indicative of a degree of relation not less than said second predetermined value.

25. (New) The document categorizing method of claim 24, wherein:

said first naming convention includes a concatenation of at least a name segment of each of said constituent evaluated clusters with a first delimiter inserted between the concatenated name segments; and

said second naming convention includes a concatenation of at least a name segment of each of said constituent evaluated clusters with a second delimiter, different from said first delimiter, inserted between the concatenated name segments.

26. (New) The document categorizing method of claim 25, wherein said second delimiter is a blank space.

27. (New) The document categorizing method of claim 25, wherein the full name of said constituent evaluated clusters are concatenated in said first and second naming conventions.

28. (New) The document categorizing method of claim 24, wherein said new combined cluster constitutes a cluster combination, said method further comprising:

determining a degree of relation between a previously uncombined cluster within said plurality of said clusters with said cluster combination by evaluating their similarity based on the documents included in said uncombined cluster and said cluster combination;

merging the evaluated uncombined cluster and the evaluated cluster combination into a newer combined cluster when their degree of relation is determined to be not less than said predetermined first value;

assigning a newer name to said newer combined cluster based on the degree of relation between its constituent evaluated previously uncombined cluster and evaluated cluster combination, wherein if their degree of relation is less than said second predetermined value, the newer name assigned to said newer combined cluster conforms to a third naming convention, and wherein if their degree of relation is not less than said second predetermined value, the newer name assigned to said newer combined cluster conforms to a fourth naming convention.

29. (New) The document categorizing method of claim 28, wherein:

said third naming convention includes a concatenation of a first part including at least a name segment of the constituent previously uncombined

cluster and a second part including the full name of the constituent combined cluster enclosed within parenthesis, and includes said first delimiter inserted between the concatenated first and second parts; and

 said fourth naming convention includes a concatenation of a first part including at least a name segment of the constituent previously uncombed cluster and a second part including the full name of the constituent combined cluster enclosed within parenthesis, and includes said second delimiter inserter between the concatenated first and second parts.

30. (New) The document categorizing method of claim 24, wherein said new combined cluster constitutes a cluster combination, said method further comprising:

 obtaining a plurality of said cluster combinations, each cluster combination having a distinctive name;

 determining a degree of relation between at least two cluster combinations by evaluating the similarity between the evaluated cluster combinations based on the documents included in the respective evaluated cluster combinations;

 merging the evaluated cluster combinations into a new combined cluster combination when their degree of relation is determined to be not less than said predetermined first value;

 assigning a new name to said new combined cluster combination based on the degree of relation between its constituent cluster combinations, wherein if the degree of relation of its constituent cluster combinations is less than said second predetermined value, the new name assigned to said new cluster combination conforms to a fifth naming convention indicative of a degree of relation between said first and second predetermined values, and wherein if the degree of relation of its constituent cluster combinations is not less than said second predetermined value, the new name assigned to said new combined cluster combination conforms to a sixth naming convention indicative of a degree of relation not less than said second predetermined value.

31. (New) The document categorizing method of claim 30, wherein:

said fifth naming convention includes a concatenation of the full name of each evaluated cluster combination, with each full name enclosed within parenthesis and separated by said first delimiter; and

 said sixth naming convention includes a concatenation of the full name of each evaluated cluster combination, with each full name enclosed within parenthesis and separated by said second delimiter.